EXPRESS MAIL LABEL NO. EU212806474US
Date of Deposit and Filing: November 6, 2003

COMPOSITIONS AND METHODS FOR THE MODIFICATION OF GENE EXPRESSION

Cross-Reference to Related Applications

This application claims priority to: U.S. Patent Application No. 10/291,447, filed November 8, 2002, which is incorporated by reference herein in its entirety, which claims priority to U.S. Provisional Patent Application 60/345,397, filed November 9, 2001; priority to U.S. Provisional Patent Application No.60/425, 087, filed November 8, 2002, which is incorporated herein by reference; priority to U.S. Patent Application No. 10/137,036, filed April 30, 2002, which claims priority to U.S. Patent Application No. 09/276,599, filed March 25, 1999 (now U.S. Patent No. 6, 380,459) and priority to U.S. Patent Application No. 09/724,624, filed November 28, 2000 (now abandoned), which is a CIP of U.S. Patent Application No. 09/598,401, filed June 20, 2000 (now U.S. Patent No. 6,596,925), which claims priority to International Patent Application No. PCT/NZ00/00018, filed February 24, 2000 and to U.S. Provisional Patent Application No. 60/146,591, filed July 30, 1999, and is a CIP of U.S. Patent Application No. 09/276,599, filed March 25, 1999 (now U.S. Patent No. 6, 380,459).

Technical Field of the Invention

This invention relates to the regulation of polynucleotide transcription and/or expression. More specifically, this invention relates to polynucleotide regulatory sequences isolated from *Eucalyptus grandis* that are capable of initiating and driving the transcription of polynucleotides in plant vascular tissues, and the use of such regulatory sequences in the modification of transcription of endogenous and/or heterologous polynucleotides involved in wood formation.

Background of the Invention

Gene expression is regulated, in part, by the cellular processes involved in transcription. During transcription, a single-stranded RNA complementary to the DNA sequence to be transcribed is formed by the action of RNA polymerases. Transcription is initiated by the binding of RNA polymerase to characteristic recognition sequences in the promoter region of the gene. As used herein, the term "promoter" refers to the 5' non-coding region of a eukaryotic gene that is involved in transcription initiation and regulation, which generally comprises between 100 and 1000 (or more) nucleotides upstream of the transcription start site. Generally, the promoter includes the transcription start site, and farther upstream from the transcription start site, the initiator region, comprising characteristic sequence motifs involved in binding RNA polymerase and initiating transcription, and *cis*-acting transcription control elements (also referred to as "promoter-proximal elements") which extend several hundred bases upstream of the transcriptional start site and interact with *trans*-acting protein factors to regulate transcription. Such *cis*-acting control elements may be cell-or-tissue specific and may determine the responsiveness of transcription of the particular gene associated with the promoter to hormones and other endogenous signals. Other *cis*-acting control elements may affect the strength of the promoter (e.g., enhancers) or efficiency of transcription (e.g., the 5' untranslated sequence downstream of the start site). Enhancers can occur upstream or downstream from the initiation site.

Tissue-specific promoters are particularly advantageous for use in transgenic modification of plants where spatial localization and/or developmental timing of gene expression is important, or where constitutive expression would be detrimental to the development and physiological function of the transgenically modified plant.

There is a continuing need for promoters that can be activated specifically in tissues involved in xylogenesis and primary and secondary xylem. Such promoters can be used to selectively modulate the expression of genes involved in secondary cell wall formation in plants, for example, by eliminating or reducing lignification (and increasing cellulose deposition) in secondary xylem, increasing the volume of particular secondary cell wall layers, and controlling the sites and levels of lignification and cellulose deposition.

Summary of the Invention

Briefly, isolated polynucleotide regulatory sequences from eucalyptus that are involved in the regulation of caffeic acid O-methyltransferase (cOMT) gene expression are disclosed, together with methods for the use of such polynucleotide regulatory regions in the modification of expression of endogenous and/or heterologous polynucleotides in transgenic plants. The invention encompasses recombinant promoters comprised of one or more motifs in the inventive promoter sequences and having new or improved activities.

In a first aspect, the present invention provides an isolated polynucleotide sequence comprising a vascular tissue-specific promoter of the *E. grandis* cOMT gene, and functional promoter fragments thereof. In a preferred embodiment, the polynucleotide sequence is selected from the group consisting of:

(a) the sequences recited in SEQ ID NO: 12 and SEQ ID 113, nucleotides 1019-1643, and their complements;

(b) reverse complements of the sequences recited in (a);

(c) reverse sequences of the sequences recited in (a);

(d) sequences having at least 75% identity to a sequence recited in (a);

(e) sequences having at least 90% identity to a sequence recited in (a);

(f) a polynucleotide sequence that hybridizes to a polynucleotide sequence of (a) above under stringent conditions; and

(g) a polynucleotide comprising a 20-mer, a 40-mer, a 60-mer, an 80-mer, a 100-mer, a 120-mer, a 150-mer, a 180-mer, a 220-mer, a 250-mer, a 300-mer, 400-mer, 500-mer or 600-mer of a sequence recited in (a) or (e) above.

In another aspect, the present invention provides genetic constructs comprising a polynucleotide of the present invention, either alone, or in combination with one or more additional polynucleotides of the present invention, or in combination with one or more known polynucleotides, together with cells and target organisms comprising such constructs. The genetic construct may comprise for example, a sequence identified herein as SEQ ID NO. 60, which comprises a cOMT promoter functionally linked to a cOMT coding sequence. This construct may be used to regulate the biosynthesis of monolignols and lignin in plants.

In one embodiment, the genetic constructs comprise, in the 5'-3' direction, a polynucleotide promoter sequence of the present invention, a polynucleotide to be transcribed, and a gene termination sequence. The polynucleotide to be transcribed may comprise an open reading frame of a polynucleotide that encodes a polypeptide of interest and/ or a non-coding, or untranslated, region of a polynucleotide of interest. The open reading frame may be orientated in either a sense or antisense direction. Preferably, the gene termination sequence is functional in a host plant. Most preferably, the gene termination sequence is that of the gene of interest, but others generally used in the art, such as the *Agrobacterium tumefaciens* nopaline synthase terminator may be usefully employed in the present invention. The genetic construct may further include a marker for the identification of transformed cells.

In one embodiment, the genetic construct is used for transcriptional silencing of a gene of interest. For example, the construct may comprise an inverted repeat of a cOMT promoter sequence or promoter fragment of the present invention driven by an unrelated promoter, such as an inducible or constitutive promoter.

In another embodiment, the genetic construct is designed to downregulate the expression of a gene of interest, for example, by posttranscriptional silencing, by antisense suppression, or by cosuppression.

In a further aspect, transgenic plant cells comprising the genetic constructs of the present invention are provided, together with organisms, such as plants, comprising such transgenic cells, and fruits, seeds and other products, derivatives, or progeny of such plants. Propagules of the inventive transgenic plants are included in the present invention. As used herein, the word "propagule" means any part of a plant that may be used in reproduction or propagation, sexual or asexual, including cuttings.

Plant varieties, particularly registerable plant varieties according to Plant Breeders' Rights, may be excluded from the present invention. A plant need not be considered a "plant variety" simply because it contains stably within its genome a transgene, introduced into a cell of the plant or an ancestor thereof.

In yet another aspect, methods for modifying gene expression in a target organism, such as a plant, are provided, such methods including stably incorporating into the genome of the organism a genetic construct of the present invention. In a preferred embodiment, the

target organism is a plant, more preferably a woody plant such as poplar, eucalyptus and pine species, sugarcane, forage grasses and *Salix* spp., most preferably from the group consisting of *Eucalyptus grandis* and *Pinus radiata*.

In another aspect, methods for producing a target organism, such as a plant, having modified polypeptide expression are provided, such methods comprising transforming a plant cell with a genetic construct of the present invention to provide a transgenic cell, and cultivating the transgenic cell under conditions conducive to regeneration and mature plant growth.

In other aspects, methods for identifying a gene responsible for a desired function or phenotype are provided, the methods comprising transforming a plant cell with a genetic construct comprising a polynucleotide promoter sequence of the present invention operably linked to a polynucleotide to be tested, cultivating the plant cell under conditions conducive to regeneration and mature plant growth to provide a transgenic plant; and comparing the phenotype of the transgenic plant with the phenotype of non-transformed, or wild-type, plants.

The above-mentioned and additional features of the present invention and the manner of obtaining them will become apparent, and the invention will be best understood by reference to the following more detailed description. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

Brief Description of the Drawings

Fig. 1. Nucleotide sequence of the Eucalyptus grandis cOMT promoter (SEQ ID NO. 12) showing the motifs (underlined) located within the sequence and the putative TATA box (grey box).

Fig. 2. Nucleotide sequence of the Eucalyptus grandis cOMT gene and promoter (SEQ ID NO. 113, nucleotides 1019-1643). The promoter region is in bold.

Fig. 3. Comparison of the promoter activity of a sequence comprising the full length 5' UTR of SEQ ID 113 (nucleotides 1-1643; cOMT 1700) with the promoter activity of SEQ ID NO: 12 (cOMT 667) in the Zinnia TE cell based assay system. pART9 is the empty

vector control. The fluorescence is represented as arbitrary fluorescence units (FU) per microgram protein per minute.

Fig. 4. Vascular-specific expression of an *E. grandis* OMT Promoter-GUS construct in transformed *Nicotiana benthamiana*. The GUS sequence was expressed under the control of nucleotides 1-1643 of SEQ ID NO. 113. Paraffin-embedded sections of the stem were stained for GUS and counterstained with Safronin O to highlight the GUS staining. Staining was observed in differentiating cambial cells and xylem.

Fig. 5. Vascular-specific expression of an *E. grandis* OMT Promoter-GUS construct in transformed *Nicotiana benthamiana*. The GUS sequence was expressed under the control of SEQ ID NO. 12. GUS staining is shown in xylem of the stem base (top left panel), midstem(top right panel) and roots (bottom panel).

Detailed Description

The present invention provides vascular tissue-specific isolated polynucleotide regulatory regions of the *Eucalyptus grandis* cOMT gene which may be used for the modification of plant phenotypes. As discussed above, promoters are components of the cellular "transcription apparatus" and are involved in the regulation of gene expression. Both tissue- and temporal-specific gene expression patterns have been shown to be initiated and controlled by promoters during the natural development of a plant. The isolated polynucleotide promoter sequences of the present invention may thus be employed in the modification of growth and development of plants, in particular, to selectively modulate the expression of genes involved in secondary cell wall formation in plants, such as those disclosed in U.S. Patent Application No. 10/198,232, filed July 17, 2002. For example, the promoter sequences of the present invention may be used for eliminating or reducing lignification (and increasing cellulose deposition) in secondary xylem, increasing the thickness of the secondary cell wall, and controlling the sites and levels of lignification and cellulose deposition in a plant.

Using the methods and materials of the present invention, the amount of a specific polypeptide of interest may be increased or reduced by incorporating additional copies of genes, or coding sequences, encoding the polypeptide, operably linked to an inventive promoter sequence, into the genome of a target organism, such as a plant. Similarly, an

increase or decrease in the amount of the polypeptide may be obtained by transforming the target plant with antisense copies or direct repeats or inverted repeats of such genes.

In one of its aspects, the present invention provides an isolated polynucleotide sequence comprising a vascular tissue-specific promoter of the E. grandis cOMT gene, and functional promoter fragments thereof. In a preferred embodiment, the polynucleotide sequence is selected from the group consisting of:

(a) the sequences recited in SEQ ID NO: 12 and SEQ ID 113, nucleotides 1019-1643, and their complements;

(b) reverse complements of the sequences recited in (a);

(c) reverse sequences of the sequences recited in (a);

(d) sequences having at least 75% identity to a sequence recited in (a);

(e) sequences having at least 90% identity to a sequence recited in (a);

(f) a polynucleotide sequence that hybridizes to a polynucleotide sequence of (a) above under stringent conditions; and

(g) a polynucleotide comprising a 20-mer, a 40-mer, a 60-mer, an 80-mer, a 100-mer, a 120-mer, a 150-mer, a 180-mer, a 220-mer, a 250-mer, a 300-mer, 400-mer, 500-mer or 600-mer of a sequence recited in (a) or (e) above.

In another embodiment, the polynucleotide sequence comprises a sequence recited in SEQ ID NO. 60 and its complement, reverse sequence, reverse complement, and sequences having at least 90% identity to these sequences.

It should be understood that the term "percent identity", as used herein, and method of calculating percent identity, are disclosed in the specification.

Polynucleotide sequences comprising a *E. grandis* cOMT promoter are shown in Figures 1 and 2, and in SEQ ID NO. 12, 60 and 113 (nucleotides 1019-1643) of the Sequence Listing.

The motifs shown in Figure 1 are considered to be putative *cis*-elements for the *E. grandis* cOMT promoters based on their similarity to known vascular specific factor-like and AC rich elements in other plant gene promoters. The sequence recited in SEQ ID NO. 12, nucleotides 1-661, is 98.9% identical to the promoter comprising sequence of SEQ ID NO. 113 (nucleotides 1019-1676). Polynucleotides comprising sequences that differ from the polynucleotide sequence recited in SEQ ID NO: 12, its complement, reverse complement or

reverse sequence, as a result of deletions and/or insertions totaling less than 10% of the total sequence length are also contemplated by and encompassed within the present invention. In certain embodiments, variants of the inventive polynucleotides possess biological activities that are the same or similar to those of the inventive polynucleotides. Such variant polynucleotides function as promoter sequences and are thus capable of modifying gene expression in a plant.

Functional fragments of SEQ ID NO. 12 and their variants having at least 90% identity are intended to be encompassed by the present invention. Functional promoter fragments can be identified by those skilled in the art using conventional deletion analysis methods and the functional assay methods described in Examples 2 and 3.

The promoter activity of the 5' noncoding region of SEQ ID NO. 113, nucleotides 1-1643 (cOMT 1700) and of SEQ ID NO: 12 (cOMT 667) in a cell based tracheary element (TE)- forming system was demonstrated by transfecting *Zinnia elegans* mesophyll cells with promoter-GUS constructs and comparing GUS expression under inducing and noninducing conditions (Example 2). As is shown in Figure 3, the expression of GUS driven by each of the above-identified sequences was much greater under TE inducing conditions than under maintenance conditions. cOMT 667 was at least as active as cOMT 1700. This experiment supported the identification of the cOMT sequence as a xylem-specific promoter.

Further experiments were performed by *in planta* analysis of the above promoter-GUS constructs in tobacco plants, as described in Example 3. Figures 4 and 5 show GUS stained sections of transformed tobacco plants. These experiments confirmed the vascular tissue-specific activity of the promoter.

The term "polynucleotide(s)," as used herein, means a single or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases and includes DNA and corresponding RNA molecules, including HnRNA and mRNA molecules, both sense and anti-sense strands, and comprehends cDNA, genomic DNA and recombinant DNA, as well as wholly or partially synthesized polynucleotides. An HnRNA molecule contains introns and corresponds to a DNA molecule in a generally one-to-one manner. An mRNA molecule corresponds to an HnRNA and DNA molecule from which the introns have been excised. A polynucleotide may consist of an entire gene, or any portion thereof. Operable anti-sense polynucleotides may comprise a fragment of the corresponding polynucleotide, and the

definition of "polynucleotide" therefore includes all such operable anti-sense fragments. Anti-sense polynucleotides and techniques involving anti-sense polynucleotides are well known in the art and are described, for example, in Robinson-Benion *et al.*, "Antisense techniques," *Methods in Enzymol.* 254:363-375, 1995; and Kawasaki *et al.*, in *Artific. Organs* 20:836-848, 1996.

All of the polynucleotides described herein are isolated and purified, as those terms are commonly used in the art. Preferably, the polynucleotides are at least about 80% pure, more preferably at least about 90% pure, and most preferably at least about 99% pure.

The definition of the terms "complement", "reverse complement" and "reverse sequence", as used herein, is best illustrated by the following example. For the sequence 5' AGGACC 3', the complement, reverse complement and reverse sequence are as follows:

Complement	3' TCCTGG 5'
Reverse complement	3' GGTCCT 5'
Reverse sequence	5' CCAGGA 3'

As used herein, the term "variant" comprehends nucleotide sequences different from the specifically identified sequences, wherein one or more nucleotides is deleted, substituted, or added. Variants may be naturally occurring allelic variants, or non-naturally occurring variants. Variant sequences preferably exhibit at least 50%, more preferably at least 75%, and most preferably at least 90% identity to a sequence of the present invention. The percentage identity is determined by aligning the two sequences to be compared as described below, determining the number of identical nucleotides in the aligned portion, dividing that number by the total number of residues in the inventive (queried) sequence, and multiplying the result by 100.

Polynucleotide sequences may be aligned, and percentage of identical residues in a specified region may be determined against other polynucleotide sequences, using computer algorithms that are publicly available. Two exemplary algorithms for aligning and identifying the similarity of polynucleotide sequences are the BLASTN and FASTA algorithms. The BLASTN algorithm Version 2.0.11 [Jan-20-2000], or later versions, set to the default parameters described in the documentation and distributed with the algorithm, are preferred for use in the determination of polynucleotide variants according to the present

invention. The use of the BLAST family of algorithms, including BLASTN, BLASTP, and BLASTX, is described in the publication of Altschul *et al.*, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Res.* 25: 3389-3402, 1997. The BLASTN software is available on the NCBI anonymous FTP server (ftp://ncbi.nlm.nih.gov) under /blast/executables/ and is available from the National Center for Biotechnology Information (NCBI), National Library of Medicine, Building 38A, Room 8N805, Bethesda, MD 20894 USA.

The FASTA software package is available from the University of Virginia (University of Virginia, PO Box 9025, Charlottesville, VA 22906-9025). Version 2.0u4, February 1996, or later versions, set to the default parameters described in the documentation and distributed with the algorithm, may be used in the determination of variants according to the present invention. The use of the FASTA algorithm is described in Pearson and Lipman, "Improved Tools for Biological Sequence Analysis," *Proc. Natl. Acad. Sci. USA* 85:2444-2448, 1988; and Pearson, "Rapid and Sensitive Sequence Comparison with FASTP and FASTA," *Methods in Enzymol.* 183:63-98, 1990.

The following running parameters are preferred for determination of alignments and similarities using BLASTN that contribute to the E values and percentage identity for polynucleotide sequences: Unix running command: blastall -p blastn -d embldb -e 10 –G0 -FF -E0 -r 1 -v 30 -b 30 -i queryseq –o results; the parameters are: -p Program Name [String]; -d Database [String]; -e Expectation value (E) [Real]; -G Cost to open a gap (zero invokes default behavior) [Integer]; -FF low complexity filter; -E Cost to extend a gap (zero invokes default behavior) [Integer]; -r Reward for a nucleotide match (BLASTN only) [Integer]; -v Number of one-line descriptions (V) [Integer]; -b Number of alignments to show (B) [Integer]; -i Query File [File In]; and -o BLAST report Output File [File Out] Optional.

The "hits" to one or more database sequences by a queried sequence produced by BLASTN and FASTA or a similar algorithm, align and identify similar portions of sequences. The hits are arranged in order of the degree of similarity and the length of sequence overlap. Hits to a database sequence generally represent an overlap over only a fraction of the sequence length of the queried sequence.

The BLASTN and FASTA algorithms also produce "Expect" values for alignments. The Expect value (E) indicates the number of hits one can "expect" to see over a certain number of contiguous sequences by chance when searching a database of a certain size. The Expect value is used as a significance threshold for determining whether the hit to a database, such as the preferred EMBL database, indicates true similarity. For example, an E value of 0.1 assigned to a polynucleotide hit is interpreted as meaning that in a database of the size of the EMBL database, one might expect to see 0.1 matches over the aligned portion of the sequence with a similar score simply by chance. By this criterion, the aligned and matched portions of the polynucleotide sequences then have a probability of 90% of being the same. For sequences having an E value of 0.01 or less over aligned and matched portions, the probability of finding a match by chance in the EMBL database is 1% or less using the BLASTN or FASTA algorithm.

According to one embodiment, "variant" polynucleotides, with reference to each of the polynucleotides of the present invention, preferably comprise sequences having the same number or fewer bases than each of the polynucleotides of the present invention and producing an E value of 0.01 or less when compared to the polynucleotide of the present invention. That is, a variant polynucleotide or polypeptide is any sequence that has at least a 99% probability of being the same as the polynucleotide or polypeptide of the present invention, measured as having an E value of 0.01 or less using the BLASTN or FASTA algorithms set at parameters described above. According to a preferred embodiment, a variant polynucleotide is a sequence having the same number or fewer nucleic acids than a polynucleotide of the present invention that has at least a 99% probability of being the same as the polynucleotide of the present invention, measured as having an E value of 0.01 or less using the BLASTN or FASTA algorithms set at parameters described above.

Alternatively, variant polynucleotides of the present invention hybridize to the polynucleotide sequences recited in SEQ ID NO. 12 and SEQ ID NO. 113, nucleotides 1019-1643, or complements, reverse sequences, or reverse complements of those sequences under stringent conditions. An example of "stringent conditions" is prewashing in a solution of 6X SSC, 0.2% SDS; hybridizing at 65°C, 6X SSC, 0.2% SDS overnight; followed by two washes of 30 minutes each in 1X SSC, 0.1% SDS at 65° C and two washes of 30 minutes

each in 0.2X SSC, 0.1% SDS at 65°C. See discussion below relating to substantial complementarity of nucleotide sequences.

The present invention also encompasses polynucleotides comprising sequences that differ from the polynucleotide sequence recited in SEQ ID NO: 12, or SEQ ID NO. 113, nucleotides 1019-1643, or complements, reverse complements or reverse sequences thereof, as a result of deletions and/or insertions totaling less than 10% of the total sequence length are also contemplated by and encompassed within the present invention. In certain embodiments, variants of the inventive polynucleotides possess biological activities that are the same or similar to those of the inventive polynucleotides. Such variant polynucleotides function as promoter sequences and are thus capable of modifying gene expression in a plant.

The polynucleotides of the present invention may be isolated from various libraries, or may be synthesized using techniques that are well known in the art. The polynucleotides may be synthesized, for example, using automated oligonucleotide synthesizers (*e.g.*, Beckman Oligo 1000M DNA Synthesizer) to obtain polynucleotide segments of up to 50 or more nucleic acids. A plurality of such polynucleotide segments may then be ligated using standard DNA manipulation techniques that are well known in the art of molecular biology. One conventional and exemplary polynucleotide synthesis technique involves synthesis of a single stranded polynucleotide segment having, for example, 80 nucleic acids, and hybridizing that segment to a synthesized complementary 85 nucleic acid segment to produce a 5-nucleotide overhang. The next segment may then be synthesized in a similar fashion, with a 5-nucleotide overhang on the opposite strand. The "sticky" ends ensure proper ligation when the two portions are hybridized. In this way, a complete polynucleotide of the present invention may be synthesized entirely *in vitro*.

Polynucleotides of the present invention also comprehend polynucleotides comprising at least a specified number of contiguous residues (x-mers) of any of the polynucleotides identified as SEQ ID NO: 12 or SEQ ID NO. 113, nucleotides 1019-1643, complements, reverse sequences, and reverse complements of such sequences, and their variants. As used herein, the term "x-mer," with reference to a specific value of "x," refers to a sequence comprising at least a specified number ("x") of contiguous residues of any of the polynucleotides identified as SEQ ID NO: 12 or SEQ ID NO. 113, nucleotides 1019-1643. According to preferred embodiments, the value of x is preferably at least 20, more preferably

at least 40, more preferably yet at least 60, and most preferably at least 80. Thus, polynucleotides of the present invention comprise a 20-mer, a 40-mer, a 60-mer, an 80-mer, a 100-mer, a 120-mer, a 150-mer, a 180-mer, a 220-mer, a 250-mer, a 300-mer, 400-mer, 500-mer or 600-mer of a polynucleotide identified as SEQ ID NO: 12 or SEQ ID NO. 113, nucleotides 1019-1643 and variants thereof.

As noted above, the inventive polynucleotide promoter sequences may be employed in genetic constructs to drive transcription and/or expression of a polynucleotide of interest. The technology involved in making expression constructs comprising the open reading frames and suitable promoters, initiators, terminators, etc., is well known in the art and can be practiced by those of ordinary skill in the art without undue experimentation. The constructs may be introduced into a host cell to express the polypeptide encoded by the open reading frame. Suitable host cells may include various prokaryotic and eukaryotic cells, including plant cells, bacterial cells, algae and the like. Procedures for transforming cells are also well known in the art and can be practiced by those of ordinary skill in the art without undue experimentation.

The polynucleotide sequence of interest may be endogenous or heterologous to the organism to be transformed, for example, a plant. The inventive genetic constructs may thus be employed to modulate levels of transcription and/or expression of a polynucleotide, for example, a gene that is present in the wild-type plant, or may be employed to provide transcription and/or expression of a polynucleotide sequence that is not normally found in the wild-type plant.

In certain embodiments, the polynucleotide of interest comprises an open reading frame that encodes a target polypeptide. Open reading frames may be identified using techniques that are well known in the art. These techniques include, for example, analysis for the location of known start and stop codons, most likely reading frame identification based on codon frequencies, etc. Suitable tools and software for ORF analysis include, for example, "GeneWise", available from the Sanger Center, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 ISA, United Kingdom; "Diogenes", available from Computational Biology Centers, University of Minnesota, Academic Health Center, UMHG Box 43, Minneapolis MN 55455 and "GRAIL", available from the Informatics Group, Oak Ridge National Laboratories, Oak Ridge, TN. The open reading frame is inserted in the

genetic construct in either a sense or antisense orientation, such that transformation of a target plant with the genetic construct will lead to a change in the amount of polypeptide compared to the wild-type plant. Transformation with a genetic construct comprising an open reading frame in a sense orientation will generally result in over-expression of the selected polypeptide, while transformation with a genetic construct comprising an open reading frame in an antisense orientation will generally result in reduced expression of the selected polypeptide. A population of plants transformed with a genetic construct comprising an open reading frame in either a sense or antisense orientation may be screened for increased or reduced expression of the polypeptide in question using techniques well known to those of skill in the art, and plants having the desired phenotypes may thus be isolated.

Alternatively, expression of a target polypeptide may be inhibited by inserting a portion of the open reading frame, in either sense or antisense orientation, in the genetic construct. Such portions need not be full-length but preferably comprise at least 25 and more preferably at least 50 residues of the open reading frame. A much longer portion or even the full length DNA corresponding to the complete open reading frame may be employed. The portion of the open reading frame does not need to be precisely the same as the endogenous sequence, provided that there is sufficient sequence similarity to achieve inhibition of the target gene. Thus a sequence derived from one species may be used to inhibit expression of a gene in a different species.

In one embodiment, RNAi expression constructs comprising partial or full-length coding sequences homologous to endogenous genes are used to downregulate the expression of these genes, for example, genes involved in biosynthesis of lignin. Representative examples of RNAi silencing methods can be found in, but are not limited to, the following patent applications and publications: PCT Applications WO 99/49029, WO 98/36083, WO 99/15682, WO 98/53083, WO 99/53050, WO 00/49035, WO 01/77350, WO01/94603, WO02/00894, WO01/75164, and WO01/68836; Waterhouse *et al.*, *Proc. Natl. Acad. Sci. USA* 95: 13959-13964 (1998).

In further embodiments, the inventive genetic constructs comprise a polynucleotide including an untranslated, or non-coding, region of a gene coding for a target polypeptide, or a polynucleotide complementary to such an untranslated region. Examples of untranslated

regions which may be usefully employed in such constructs include introns and 5'-untranslated leader sequences. Transformation of a target plant with such a genetic construct may lead to a reduction in the amount of the polypeptide expressed in the plant by the process of cosuppression, in a manner similar to that discussed, for example, by Napoli *et al.*, *Plant Cell* 2:279-290, 1990 and de Carvalho Niebel *et al.*, *Plant Cell* 7:347-358, 1995.

The present invention also contemplates using a genetic construct to produce gene knockouts by transcriptional gene silencing, for example, a construct comprising an inverted repeat of a promoter sequence or a promoter fragment of the present invention under the control of a different promoter (see, e.g., Mette *et al.*, *EMBO J.* 19: 5194-5201, 2000).

In one embodiment of the invention, the polynucleotide of interest, such as a coding sequence in sense or antisense orientation, is operably linked to a polynucleotide promoter sequence of the present invention such that the promoter directs transcription of the coding sequence in vascular tissues, preferably xylem or a tissue involved in xylogenesis (e.g., cambium initials of vascular cambium). The polynucleotide promoter sequence is generally positioned at the 5' end of the polynucleotide to be transcribed. The promoter can be employed to modulate gene transcription during xylem development in a transformed plant.

The properties of the cOMT promoter, e.g., its strength and inducibility or hormone responsiveness, can be modified by deletion, insertion, rearrangement or mutation of cis-acting elements to obtain new or improved properties (e.g., activation by external stimuli, such as light, heat, anaerobic stress, alteration in nutrient conditions, chemicals and the like).

The inventive genetic constructs further comprise a gene termination sequence which is located 3' to the polynucleotide of interest. A variety of gene termination sequences which may be usefully employed in the genetic constructs of the present invention are well known in the art. One example of such a gene termination sequence is the 3' end of the *Agrobacterium tumefaciens* nopaline synthase gene. The gene termination sequence may be endogenous to the target plant or may be exogenous, provided the promoter is functional in the target plant. For example, the termination sequence may be obtained from other plant species, plant viruses, bacterial plasmids and the like.

The genetic constructs of the present invention may also contain a selection marker that is effective in cells of the target organism, such as a plant, to allow for the detection of transformed cells containing the inventive construct. Such markers, which are well known in

the art, typically confer resistance to one or more toxins. One example of such a marker is the NPTII gene whose expression results in resistance to kanamycin or hygromycin, antibiotics which are usually toxic to plant cells at a moderate concentration (Rogers *et al.*, *in* Weissbach A and H, eds. *Methods for Plant Molecular Biology*, Academic Press Inc.: San Diego, CA, 1988). Transformed cells can thus be identified by their ability to grow in media containing the antibiotic in question. Alternatively, the presence of the desired construct in transformed cells can be determined by means of other techniques well known in the art, such as Southern and Western blots.

Techniques for operatively linking the components of the inventive genetic constructs are well known in the art and include the use of synthetic linkers containing one or more restriction endonuclease sites as described, for example, by Sambrook *et al.*, *Molecular cloning: a laboratory manual*, CSHL Press: Cold Spring Harbor, NY, 1989. The genetic construct of the present invention may be linked to a vector having at least one replication system, for example *E. coli*, whereby after each manipulation, the resulting construct can be cloned and sequenced and the correctness of the manipulation determined.

The genetic constructs of the present invention may be used to transform a variety of target organisms including, but not limited to, plants. Plants which may be transformed using the inventive constructs include both monocotyledonous angiosperms (*e.g.*, grasses, corn, grains, oat, wheat and barley) and dicotyledonous angiosperms (*e.g.*, *Arabidopsis*, tobacco, legumes, alfalfa, oaks, eucalyptus, maple), and Gymnosperms (*e.g.*, Scots pine; *see* Aronen, *Finnish Forest Res. Papers*, Vol. 595, 1996), white spruce (Ellis *et al.*, *Biotechnology* 11:84-89, 1993), and larch (Huang *et al.*, *In Vitro Cell* 27:201-207, 1991). In a preferred embodiment, the inventive genetic constructs are employed to transform grasses and woody plants. Woody plants are herein defined as a tree or shrub whose stem lives for a number of years and increases in diameter each year by the addition of woody tissue. Preferably the woody plant is selected from the group consisting of eucalyptus and pine species, most preferably from the group consisting of *Eucalyptus grandis* and *Pinus radiata*. Other preferred species include Poplar, sugarcane, forage grasses and *Salix* spp. Other species which may be usefully transformed with the genetic constructs of the present invention include, but are not limited to: pines such as *Pinus banksiana, Pinus brutia, Pinus caribaea, Pinus clausa, Pinus contorta, Pinus coulteri, Pinus echinata, Pinus eldarica, Pinus*

ellioti, Pinus jeffreyi, Pinus lambertiana, Pinus monticola, Pinus nigra, Pinus palustrus, Pinus pinaster, Pinus ponderosa, Pinus resinosa, Pinus rigida, Pinus serotina, Pinus strobus, Pinus sylvestris, Pinus taeda, Pinus virginiana; other gymnosperms, such as *Abies amabilis, Abies balsamea, Abies concolor, Abies grandis, Abies lasiocarpa, Abies magnifica, Abies procera, Chamaecyparis lawsoniona, Chamaecyparis nootkatensis, Chamaecyparis thyoides, Juniperus virginiana, Larix decidua, Larix laricina, Larix leptolepis, Larix occidentalis, Larix siberica, Libocedrus decurrens, Picea abies, Picea engelmanni, Picea glauca, Picea mariana, Picea pungens, Picea rubens, Picea sitchensis, Pseudotsuga menziesii, Sequoia gigantea, Sequoia sempervirens, Taxodium distichum, Tsuga canadensis, Tsuga heterophylla, Tsuga mertensiana, Thuja occidentalis, Thuja plicata*; and Eucalypts, such as *Eucalyptus alba, Eucalyptus bancroftii, Eucalyptus botyroides, Eucalyptus bridgesiana, Eucalyptus calophylla, Eucalyptus camaldulensis, Eucalyptus citriodora, Eucalyptus cladocalyx, Eucalyptus coccifera, Eucalyptus curtisii, Eucalyptus dalrympleana, Eucalyptus deglupta, Eucalyptus delagatensis, Eucalyptus diversicolor, Eucalyptus dunnii, Eucalyptus ficifolia, Eucalyptus globulus, Eucalyptus gomphocephala, Eucalyptus gunnii, Eucalyptus henryi, Eucalyptus laevopinea, Eucalyptus macarthurii, Eucalyptus macrorhyncha, Eucalyptus maculata, Eucalyptus marginata, Eucalyptus megacarpa, Eucalyptus melliodora, Eucalyptus nicholii, Eucalyptus nitens, Eucalyptus nova-anglica, Eucalyptus obliqua, Eucalyptus obtusiflora, Eucalyptus oreades, Eucalyptus pauciflora, Eucalyptus polybractea, Eucalyptus regnans, Eucalyptus resinifera, Eucalyptus robusta, Eucalyptus rudis, Eucalyptus saligna, Eucalyptus sideroxylon, Eucalyptus stuartiana, Eucalyptus tereticornis, Eucalyptus torelliana, Eucalyptus urnigera, Eucalyptus urophylla, Eucalyptus viminalis, Eucalyptus viridis, Eucalyptus wandoo* and *Eucalyptus youmanni*; and hybrids of any of these species.

Techniques for stably incorporating genetic constructs into the genome of target plants are well known in the art and include *Agrobacterium tumefaciens* mediated introduction, electroporation, protoplast fusion, injection into reproductive organs, injection into immature embryos, high velocity projectile introduction and the like. The choice of technique will depend upon the target plant to be transformed. For example, dicotyledonous plants and certain monocots and gymnosperms may be transformed by *Agrobacterium* Ti plasmid technology, as described, for example by Bevan, *Nucleic Acids Res.* 12:8711-8721, 1984. Targets for the introduction of the genetic constructs of the present invention include

tissues, such as leaf tissue, dissociated cells, protoplasts, seeds, embryos, meristematic regions; cotyledons, hypocotyls, and the like. The preferred method for transforming eucalyptus and pine is a biolistic method using pollen (see, for example, Aronen, *Finnish Forest Res. Papers*, Vol. 595, 53pp, 1996) or easily regenerable embryonic tissues.

Once the cells are transformed, cells having the inventive genetic construct incorporated in their genome may be selected by means of a marker, such as the kanamycin resistance marker discussed above. Transgenic cells may then be cultured in an appropriate medium to regenerate whole plants, using techniques well known in the art. In the case of protoplasts, the cell wall is allowed to reform under appropriate osmotic conditions. In the case of seeds or embryos, an appropriate germination or callus initiation medium is employed. For explants, an appropriate regeneration medium is used. Regeneration of plants is well established for many species. For a review of regeneration of forest trees see Dunstan *et al.*, "Somatic embryogenesis in woody plants," *in* Thorpe TA, ed., *In Vitro Embryogenesis of Plants* (*Current Plant Science and Biotechnology in Agriculture Vol. 20*), Chapter 12, pp. 471-540, 1995. Specific protocols for the regeneration of spruce are discussed by Roberts *et al.*, "Somatic embryogenesis of spruce," *in* Redenbaugh K, ed., *Synseed: applications of synthetic seed to crop improvement,* CRC Press: Chapter 23, pp. 427-449, 1993). Transformed plants having the desired phenotype may be selected using techniques well known in the art. The resulting transformed plants may be reproduced sexually or asexually, using methods well known in the art, to give successive generations of transgenic plants.

As discussed above, the production of RNA in target cells can be controlled by choice of the promoter sequence, or by selecting the number of functional copies or the site of integration of the polynucleotides incorporated into the genome of the target host. A target organism may be transformed with more than one genetic construct of the present invention, thereby modulating the activity of more than gene. Similarly, a genetic construct may be assembled containing more than one open reading frame coding for a polypeptide of interest or more than one untranslated region of a gene coding for such a polypeptide.

The isolated polynucleotides of the present invention and their variants may be used to design oligonucleotide probes and primers, for use in detecting and isolating xylem-specific promoters of genes in other plant species. Oligonucleotide probes and primers designed using the polynucleotides of the present invention may also be used in connection

with various microarray technologies, including the microarray technology of Affymetrix (Santa Clara, CA).

The term "vascular tissue-specific promoter" refers to a promoter expressed in vascular tissue of a plant, for example, xylem, phloem or vascular cambium which gives rise to these tissues. For purposes of the present application, vascular tissue-specific refers to a promoter which is preferentially expressed in xylem, phloem or cambium.

As used herein, the term "oligonucleotide" refers to a relatively short segment of a polynucleotide sequence, generally comprising between 6 and 60 nucleotides, and comprehends both probes for use in hybridization assays and primers for use in the amplification of DNA by polymerase chain reaction.

An oligonucleotide probe or primer is described as "corresponding to" a polynucleotide of the present invention, including one of the sequences set out as SEQ ID NO: 12, or SEQ ID NO. 113, nucleotides 1019-1643, or a variant thereof, if the oligonucleotide probe or primer, or its complement, is contained within one of the inventive promoter sequences, or a variant thereof. Oligonucleotide probes and primers of the present invention are substantially complementary to a polynucleotide disclosed herein.

Two single stranded sequences are said to be substantially complementary when the nucleotides of one strand, optimally aligned and compared, with the appropriate nucleotide insertions and/or deletions, pair with at least 80%, preferably at least 90% to 95% and more preferably at least 98% to 100% of the nucleotides of the other strand. Alternatively, substantial complementarity exists when a first DNA strand will selectively hybridize to a second DNA strand under stringent hybridization conditions. Stringent hybridization conditions for determining complementarity include salt conditions of less than about 1 M, more usually less than about 500 mM, and preferably less than about 200 mM. Hybridization temperatures can be as low as 5°C, but are generally greater than about 22°C, more preferably greater than about 30°C, and most preferably greater than about 37°C. Longer DNA fragments may require higher hybridization temperatures for specific hybridization. Since the stringency of hybridization may be affected by other factors such as probe composition, presence of organic solvents and extent of base mismatching, the combination of parameters is more important than the absolute measure of any one alone.

In specific embodiments, the oligonucleotide probes and/or primers comprise at least about 6 contiguous residues, more preferably at least about 10 contiguous residues, and most preferably at least about 20 contiguous residues complementary to a polynucleotide sequence of the present invention. Probes and primers of the present invention may be from about 8 to 100 base pairs in length or, preferably from about 10 to 50 base pairs in length or, more preferably from about 15 to 40 base pairs in length. The probes can be easily selected using procedures well known in the art, taking into account DNA-DNA hybridization stringencies, annealing and melting temperatures, and potential for formation of loops and other factors, which are well known in the art.. Preferred techniques for designing PCR primers are disclosed in Dieffenbach, CW and Dyksler, GS. *PCR Primer: a laboratory manual*, CSHL Press: Cold Spring Harbor, NY, 1995. A software program suitable for designing probes, and especially for designing PCR primers, is available from Premier Biosoft International, 3786 Corina Way, Palo Alto, CA 94303-4504.

A plurality of oligonucleotide probes or primers corresponding to a polynucleotide of the present invention may be provided in a kit form. Such kits generally comprise multiple DNA or oligonucleotide probes, each probe being specific for a polynucleotide sequence. Kits of the present invention may comprise one or more probes or primers corresponding to a polynucleotide of the present invention.

In one embodiment useful for high-throughput assays, the oligonucleotide probe kits of the present invention comprise multiple probes in an array format, wherein each probe is immobilized at a predefined, spatially addressable location on the surface of a solid substrate. Array formats which may be usefully employed in the present invention are disclosed, for example, in U.S. Patents No. 5,412,087 and 5,545,451; and PCT Publication No. WO 95/00450, the disclosures of which are hereby incorporated by reference.

EXAMPLES

The following examples are offered by way of illustration and not by way of limitation.

EXAMPLE 1

Isolation of an O-Methyl Transferase Promoter from *Eucalyptus grandis*

E. grandis cDNA expression libraries were constructed and screened as follows. mRNA was extracted from plant tissue using the protocol of Chang *et al.*, *Plant Molecular Biology Reporter* 11:113-116, 1993 with minor modifications. Specifically, samples were dissolved in CPC-RNAXB (100 mM Tris-Cl, pH 8,0; 25 mM EDTA; 2.0 M NaCl; 2%CTAB; 2% PVP and 0.05% Spermidine*3HCl) and extracted with chloroform:isoamyl alcohol, 24:1. mRNA was precipitated with ethanol and the total RNA preparate was purified using a Poly(A) Quik mRNA Isolation Kit (Stratagene, La Jolla, CA). A cDNA expression library was constructed from the purified mRNA by reverse transcriptase synthesis followed by insertion of the resulting cDNA clones in Lambda ZAP using a ZAP Express cDNA Synthesis Kit (Stratagene), according to the manufacturer's protocol. The resulting cDNAs were packaged using a Gigapack II Packaging Extract (Stratagene) employing 1 µl of sample DNA from the 5 µl ligation mix. Mass excision of the library was done using XL1-Blue MRF' cells and XLOLR cells (Stratagene) with ExAssist helper phage (Stratagene). The excised phagemids were diluted with NZY broth (Gibco BRL, Gaithersburg, MD) and plated out onto LB-kanamycin agar plates containing 5-bromo-4-chloro-3-indolyl-beta-D-galactosidase (X-gal) and isopropylthio-beta-galactoside (IPTG).

Of the colonies plated and picked for DNA miniprep, 99% contained an insert suitable for sequencing. Positive colonies were cultured in NZY broth with kanamycin and cDNA was purified by means of alkaline lysis and polyethylene glycol (PEG) precipitation. Agarose gel at 1% was used to screen sequencing templates for chromosomal contamination. Dye primer sequences were prepared using a Turbo Catalyst 800 machine (Perkin Elmer/Applied Biosystems Division, Foster City, CA) according to the manufacturer's protocol.

DNA sequences for positive clones was obtained using a Perkin Elmer/Applied Biosystems Division Prism 377 sequencer. cDNA clones were sequenced first from the 5' end and, in some cases, also from the 3' end. For some clones, internal sequence was obtained using subcloned fragments. Subcloning was performed using standard procedures of restriction mapping and subcloning to pBluescript II SK+ vector.

Plant polynucleotide sequences homologous to a caffeic acid O-methyl transferase (cOMT) gene, which encodes an enzyme that functions in lignin synthesis, were isolated

from a *Eucalyptus grandis* cDNA expression library. Analysis by PCR and DNA sequencing confirmed the identity of these sequences. Using the "Genome Walker" kit (Clontech, Palo Alto, CA) and gene specific primers designed from the plant polynucleotide sequence, 5'UTR sequences containing the promoter of the *E. grandis* OMT gene were isolated from genomic DNA and were extended by further sequencing. The determined nucleotide sequence is given in SEQ ID NO: 12. Figure 1 shows the cOMT promoter sequence (SEQ ID NO. 12), with putative cis-regulatory motifs. This promoter sequence was extended by further sequencing. The extended cDNA sequences are given in SEQ ID NO: 60 and 113. Figure 2 shows the full length polynucleotide sequence which includes the 5'UTR comprising the promoter sequence (in bold type), and the coding sequence.

EXAMPLE 2
Analysis of promoter activity using the Zinnia tracheary element (TE) assay

Zinnia elegans mesophyll cells were cultured in maintenance medium (FK) or TE inducing medium (FKH) as described previously (WO 03/040,403, which is incorporated herein by reference in its entirety). Protoplasts were isolated and transformed with a plasmid comprising the GUS (β-D-glucuronidase) reporter gene in frame with *E. grandis* cOMT promoter-containing sequences, as follows.

Zinnia protoplasts in 24% sucrose solution were overlaid with 1 ml of W5 solution and centrifuged at 70 x g for 10 minutes at 20° C with brakes off. Floating protoplasts were harvested and resuspended in 10 ml of W5 solution. Protoplasts were pelleted by centrifuging at 70 x g for 10 minutes at 20° C. Protoplasts were resuspended in MaMg medium (density = ~5 x 10^6 protoplasts/ml) and aliquoted into individual 15 ml tubes (300 μL: 1.5 x 10^6 protoplasts). 5 μg DNA (of each construct) and 50 μg Salmon Testes DNA was added to the protoplast suspension, mixed and incubated for 5 minutes at 20° C. 300 μl 40% PEG solution was added to each aliquot of protoplasts, mixed and incubated for 20 minutes at 20°C. 5 ml of K3/0.4M sucrose was added to each aliquot of leaf-derived transfected protoplasts or transfected protoplasts from mesophyll cells cultured in FK medium and mixed. Similarly, 5 ml of K3/0.4M sucrose+0.1ppm NAA+0.2ppm BA was added to each aliquot of transfected protoplasts from mesophyll cells cultured in FKH medium and mixed. The transfected protoplast suspensions were incubated overnight at 23°C in the dark.

Transfected *Zinnia* protoplast suspensions prepared as described above were individually harvested by adding 9.5 ml of W5 solution, mixing the contents of each tube and centrifuging at 70 x g for 10 minutes at 20° C. The bulk of the supernatant was removed by decanting and the protoplasts volume was brought up to 900 µl. From this, 300 µl of protoplasts were aliquoted into 5 ml polystyrene round-bottom tubes, re-suspended in a volume of 500 µl W5 medium and set aside for analysis of fluorescent reporter gene expression and cell viability. The protoplasts and the remaining solution were transferred to individual microtubes and pelleted by centrifugation at 420 x g for 2 minutes at 20° C. The protoplast pellet was assayed for GUS reporter gene expression as described by Jefferson, R.A., 1987, *Plant Mol. Biol. Rep.* 5, 387. GUS (MUG) assays were performed using a Wallac (Turku, Finland) Victor ^{2} 1420 Multilabel Counter. Umbelliferone was detected using a 355 nm excitation filter and a 460 nm emission filter for 1 second.

EXAMPLE 3

Determination of the activity of *E. grandis* cOMT promoter constructs in transformed *Nicotiana benthamiana*

Nicotiana benthamiana transgenic lines were transformed with *A. tumefaciens* containing constructs comprising *E. grandis* OMT promoter sequences (SEQ ID NO: 113, bases 1-1643 or SEQ ID NO: 12) operatively linked to the GUS reporter gene. Transformed plants were hand-sectioned and examined microscopically for GUS staining. The promoter sequences were cloned into plasmid pBI-101 containing a GUS reporter gene.

Agrobacterium tumefaciens transformation

Agrobacterium tumefaciens strain GV3101 was transformed with these constructs using electroporation. Electrocompetent *A. tumefaciens* cells were prepared according to the method of Walkerpeach and Velten, *Plant Mol. Biol. Man.* B1:1-19, 1994. Construct DNA (4 ng) was added to 40 µl competent *A. tumefaciens* GV3101 cells and electroporation was carried out using a BTX Electro Cell Manipulator 600 at the following settings: Mode: T 2.5 kV Resistance high voltage (HV), Set Capacitance: C (not used in HV mode), Set Resistance: R R5 (129 Ohm), Set charging voltage: S 1.44 kV, Desired field strength: 14.4 kV/cm and Desired pulse strength: t 5.0 msec. 400 µl YEP liquid media (20g/l yeast, 20

g/l peptone and 10 g/l sodium chloride) was added to the cuvette and the cells were allowed to recover for one hour at room temperature. Transformed bacteria in YEP medium were spread out on solid YEP medium containing 50 mg/l kanamycin and 50 mg/l rifampicin and incubated at 28°C for two days to allow colony growth.

Confirmation of transformation of constructs into *A. tumefaciens*

To confirm that the constructs were transformed into *A. tumefaciens*, DNA from the *A. tumefaciens* colonies from the YEP plates were isolated using standard protocols and amplified using the polymerase chain reaction (PCR) with primers designed from the pBI-101 vector sequence. The primer sequences are given in SEQ ID NOS: 128 and 129 of the Sequence Listing. PCR reactions were set up following standard protocols and 30 PCR cycles were done with extension temperature of 72°C.

Transformation of *Nicotiana benthamiana* leaf explants with *Agrobacterium*

In the laminar flow hood, 6 – 8 young, expanding leaves were removed from *N. benthamiana* plants and cut into 5 mm squares. Approximately 10 pieces from each leaf were transferred to a deep Petri dish containing ToCo medium (MS + 1 mg/l BA + 0.1 mg/l NAA). *A. tumefaciens* was grown in YEP medium containing 50 mg/l kanamycin and 50 mg/l rifampicin at 28°C for 16 hours, then centrifuged to pellet the cells. The cells were resuspended in MS liquid medium (Murashige and Skoog; Sigma, St Louis MI) without added hormone to an OD_{600} of approximately 0.8. 15 to 20 ml of the *Agrobacterium* culture (OD_{600} = 0.8) was added to each plate and soaked for 5 min with occasional gentle mixing. The leaf discs were blotted dry with sterile tissue paper, and 20 discs per plate were placed onto fresh ToCo plates with the abaxial surface uppermost. The plates were sealed and incubated under lights in ambient conditions for 2 to 3 days.

Transfer of *N. benthamiana* leaf explants co-cultivated with *A. tumefaciens* to ToSe selection medium plates.

After 3 days of co-cultivation, leaf explants were transferred from the ToCo co-cultivation plates onto ToSe selection plates (MS + 1 mg/l BA + 0.1 mg/l NAA, Timentin 200mg/l and Kanamycin 100 mg/l), sealed and incubated under lights in ambient conditions.

After 4 weeks on the selection medium, explants were cut into 4 to 8 pieces, transferred to fresh ToSe selection medium, sealed and incubated under lights in ambient conditions.

Subculture of putative transgenic *N. benthamiana* explants onto rooting medium (ToRt).

After subculture to elongation medium ToE1 (MS without hormone, Timentin 200 mg/l and Kanamycin 100 mg/l), multiple buds/shoots per callus were observed on the explants. Healthy shoots approx. 1 cm long were excised, one per callus, and transferred to rooting medium ToRt (MS without hormone, Timentin 200 mg/l and Kanamycin 100 mg/l) in tubs. Other shoots on the same callus were discarded to avoid duplicating the same transformation events. The tubs were sealed and incubated under lights in ambient conditions

After 3 to 4 weeks, the plants were transferred to soil and grown at 22°C with a 12 hour photoperiod.

GUS expression in transformed tobacco sections

Three weeks after transferring transformed plants to soil, the first side branch was sacrificed, and the following tissues were sectioned from the branch: 5-8 mm section at the base, mid and tip areas of the side branch, a leaf, root material and a floral bud. These tissues were immersed in GUS staining solution (0.5% Triton X-100, 1 mM X-GlcA-sodium salt (Duchefa, Haarlem, The Netherlands), 50 mM sodium phosphate buffer (pH 7.2), 20% methanol, pH 7.3) in an Eppendorf tube and vacuum infiltrated 3 times for 5 min. The samples were placed in a dark box, and incubated overnight at 37 °C on a shaker to allow color development. A preliminary check was done after 4 hours to detect any samples with early color development. The next day, the GUS solution was removed and the tissues destained using 70% Ethanol. The tissues were then photographed.

For photography, the sections of stem, leaf and root tissue were hand-sectioned and imaged using a microscope utilizing Metamorph software (Universal Imaging Corporation, Downingtown PA). The floral bud was sectioned down the middle, and then photographed using the Metamorph software.

Stems from 2 to 3 tobacco plants that showed good expression were collected and cut into sections of smaller than 5 mm. The tissue sections were placed in vials containing freshly-prepared fixative (25 ml 100% ethanol, 2.5 ml glacial acetic acid, 5 ml 37%

formaldehyde (Analar) and 17.5 ml water), with a vacuum applied twice for 15 min. The samples were incubated in fixative for 2 hours, a vacuum is applied for 15 min and the tissues were fixed overnight at 4°C. The fixative was removed and the tissue sections were dehydrated with a series of ethanol solutions (50% x 2, 60%, 70%, 85%), each for 30 min at room temperature. The 85% ethanol solution was removed and replaced with 95% ethanol, and the samples were incubated overnight.

The tissues were then passaged into xylene at room temperature, according to the following protocol. The 95% ethanol was removed and replaced with 100% ethanol, 60 min. This solution was removed and replaced with fresh 100% ethanol for 30 min., which was removed and replaced with 25% Xylene: 75% Ethanol, 30 min, which was removed and replaced with 50% Xylene: 50% Ethanol, 30 min, which was removed and replaced with 75% Xylene:25% Ethanol, 30 min, which was removed and replaced 3x with 100% Xylene, for 60 min each time.

The tissues were transferred to vials containing xylene and 20 paraplast chips. The vials were capped and incubated at room temperature overnight.

The vials were placed at 42°C and left until the paraplast chips had dissolved. A total of 60-80 paraffin chips were added and allowed to dissolve during the day. The vials were then incubated at 62°C overnight. During the next 48 hours, the paraplast was changed every 12 hours.

The tissues were embedded by filling an embedding cassette with liquid paraffin, placing the tissue in the correct orientation, and placing the cassette on a flat surface at 4°C overnight to allow the paraffin to harden.

Safranin O staining

Safranin O stain was prepared by dissolving 50 mg Safranin O (Raymon A Lamb Waxes and General Lab supplies, Wembley, Middlesex, UK) in a mixture of 2 ml ethanol/8 ml water. The stain was further diluted by adding 8 µl Safranin O to 1 ml of the ethanol/water mixture. Tissue was stained for 30 seconds, and the slide was dried and viewed.

Detection of GUS expression in tobacco plants at two to three months after soil transfer

Seed was collected from all plants. Two or three plants that showed expression at the 3 week stage were selected and analyzed for GUS expression. The main stem was removed from the plants, as well as any other tissue that showed initial expression. These were stained for GUS expression as described above for the 3-week samples, except that the vacuum infiltration was done twice for 30 min.

All of the eight plants that were tested with the promoter construct comprising SEQ ID NO. 12 showed vascular tissue-specific GUS expression. The expression was specifically located in the xylem cells (lignified) of the stem (base, mid and tip sections), and was not observed in the leaf or floral material. Expression was also observed in three of the roots tested. These results are comparable to those obtained with constructs comprising the 5' UTR of SEQ ID NO. 113 (nucleotides 1-1643), where expression was confined to xylem cells and differentiating cambium, but was not detected in nonvascular tissues.

While the present invention has been described with reference to the specific embodiments thereof, it should be understood by those skilled in the art that various changes may be made and equivalents may be substituted without departing from the true spirit and scope of the invention. In addition, many modifications may be made to adapt a particular situation, material, composition of matter, method, method step or steps, for use in practicing the present invention. All such modifications are intended to be within the scope of the claims appended hereto.

All of the publications, patent applications and patents cited in this application are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent application or patent was specifically and individually indicated to be incorporated by reference in its entirety.